The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t716100704

The role of working memory in the carry operation of mental arithmetic: Number and value of the carry

Ineke Imbo [a]; André Vandierendonck [a]; Stijn De Rammelaere [a]
[a] Ghent University, Ghent, Belgium

First Published on: 20 July 2006
To cite this Article: Imbo, Ineke, Vandierendonck, André and De Rammelaere, Stijn (2006) 'The role of working memory in the carry operation of mental arithmetic: Number and value of the carry', The Quarterly Journal of Experimental Psychology, 60:5, 708 - 731
To link to this article: DOI: 10.1080/17470210600762447
URL: http://dx.doi.org/10.1080/17470210600762447

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2007, 60 (5), 708–731

Psychology Press
Taylor & Francis Group

The role of working memory in the carry operation of mental arithmetic: Number and value of the carry

Ineke Imbo, André Vandierendonck, and Stijn De Rammelaere
Ghent University, Ghent, Belgium

Two experiments were conducted to investigate the role of phonological and executive working-memory components in the carry operation in mental arithmetic. We manipulated the number of carry operations, as previous research had done, but also the value that had to be carried. Results of these experiments show that in addition to the number of carry operations, the value of the carry is also an important variable determining the difficulty of arithmetical sums. Furthermore, both variables (number and value) interacted with each other in such a way that the combination of multiple carries and values of carries larger than one resulted in more difficult problems irrespective of the presence of a working-memory load. The findings with respect to working-memory load suggest that mainly the central executive is important in handling the number of carry operations as well as the value that has to be carried. The implications of the present findings for our views on mental arithmetic and its reliance on working memory are discussed.

The present article focuses on the role of working memory in handling carries in multidigit mental addition problems. This builds on previous research concerning the role of working memory in mental arithmetic with respect to both solving multidigit sums and handling the presence of carries. A recent literature review by DeStefano and LeFevre (2004) summarizes the evidence on the role of working memory in mental arithmetic. This review shows that the working-memory model of Baddeley and Hitch (1974) is the dominantly used model in this type of research. The model has no doubt the advantage that a dissociation can be made between effects due to controlled processing (the central executive) and effects due to maintenance of modality-specific information in the capacity-limited slave systems: the phonological loop and the visuo-spatial sketch pad. Therefore, and also for reasons of continuity with previous research, the present study also used this model to frame the research question.

A clear finding of the research so far is that for operations on single-digit numbers (e.g., De Rammelaere, Stuyven, & Vandierendonck, 1999, 2001; De Rammelaere & Vandierendonck,

Correspondence should be addressed to André Vandierendonck, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Gent, Belgium. E-mail: Ineke Imbo@UGent.be *or* Andre.Vandierendonck@UGent.be

Stijn De Rammelaere is currently a research consultant at Synovate.

The research reported in this article was supported by Grant No. 011D07803 of the Special Research of Ghent University to the first author and by Grant No. 10251101 of the same agency to the second author.

We are indebted to Marc Brysbaert and to two anonymous reviewers for their helpful comments on an earlier version of the present article.

http://www.psypress.com/qjep DOI:10.1080/17470210600762447

2001; Lemaire, Abdi, & Fayol, 1996; Seitz & Schumann-Hengsteler, 2002) and for operations on multidigit numbers (e.g., Fürst & Hitch, 2000; Logie, Gilhooly, & Wynn, 1994; Seitz & Schumann-Hengsteler, 2002), the central executive seems indispensable.

The role of the other working-memory components is far less clear. In one-digit arithmetic, for instance, most studies did not find a role for the phonological loop (De Rammelaere et al., 1999, 2001; De Rammelaere & Vandierendonck, 2001), whereas in multidigit arithmetic, several studies have found evidence for a role of the phonological loop in maintaining interim results (e.g., Ashcraft & Kirk, 2001; Heathcote, 1994; Logie et al., 1994; Seitz & Schumann-Hengsteler, 2002; Trbovich & LeFevre, 2003) at least for addition, while for the other operations the situation remains unclear. The visuo-spatial sketch pad may be involved in the maintenance of interim results when the participants are encouraged to use a visual problem representation (Logie et al., 1994).

The presence of carries, as in 457 + 268, is commonly considered as a factor increasing problem difficulty. To the extent that working memory is required to handle carries in dual-task experiments, an interaction between working-memory load and problem difficulty is expected. Fürst and Hitch (2000, Exp. 2) confirmed this for 0, 1, or 2 carries in the addition of two 3-digit numbers by showing that under an executive load (a variant of the Trails task) many more errors were committed on 2-carry problems (45%) than on the same problems in the articulatory suppression (15%) or control (12%) conditions. Similarly, Seitz and Schumann-Hengsteler (2002) used two-digit plus two-digit sums with or without carry and showed that the amount of interference was larger with an executive load (random generation) than with articulatory suppression such that solution of carry problems was slowed more than that of no-carry problems. In contrast, a study in which participants were required to perform a sequence of two-digit additions (Logie et al., 1994) did not observe an interaction of working memory with the presence of carry operations. As pointed out by the authors, this observation may be the result of a lack of statistical sensitivity of the experiment, but it may also suggest that working memory does not mediate carrying.

Ashcraft and Kirk (2001) required their participants to retain a two-letter or a six-letter load while solving arithmetic sums of one- or two-digit numbers. They found slower solutions of the carry problems than of the no-carry problems and also slower solution of the large versus the small letter load. In accuracy more errors were committed in the carry than in the no-carry problems, and this effect was increased under larger loads. As pointed out by DeStefano and LeFevre (2004), this finding may indicate that the phonological loop is required to maintain the carry information, but it may also "reflect the demands for central executive resources of the difficult six-letter load" (p. 371).

Other researchers addressed the role of the slave systems, and in particular the phonological loop, in carrying. Logie et al. (1994) did not observe differential effects of load on carry versus no-carry problems for either a phonological or a visuo-spatial load. In contrast, Noël, Désert, Aubrun, and Seron (2001) provided evidence for a mediation of the phonological loop in solving carry problems. They manipulated visual and phonological similarity of the digits to be added in additions of two three-digit numbers. They found an interaction of phonological similarity by problem difficulty (0 vs. 2 carries) in the reaction times. However, due to the design of the experiment with brief sequential presentation of the operands, it is not clear whether this effect was due to the higher demands placed by the carries on the phonological loop or to a combination of carry manipulation and maintenance of operands and of interim results (see also DeStefano & LeFevre, 2004).

In summary, it may be said that although there is some convincing evidence that carry manipulation is mediated by the central executive, it remains unclear whether the phonological loop plays a role in handling carry information, since some studies did while others did not find an effect of parameters related to phonological storage. A few exceptions notwithstanding (Fürst & Hitch, 2000; Noël et al., 2001), most studies compared no-carry versus one-carry problems, so that there is no clear evidence whether the number of carries in a problem is of importance. Furthermore, all studies were restricted to sums of two numbers so that the value of the carry never exceeded one (an exception must be made for Logie et al., 1994).

In order to improve our knowledge about the role of working memory in the manipulation of carries, the present study manipulated working-memory load, the number of carries in a problem and the value of the carries. In the context of a task as used by Fürst and Hitch (2000), where the problem remains visible until it is solved and where the answer is given in the order units, tens, and so on, the need to maintain interim results is limited. If there are no carries, the problem is just a concatenation of simpler mental arithmetic problems. For example, the sum 36 + 23 can be solved by decomposing this problem into simple arithmetic problems, namely 6 + 3 for the units and 3 + 2 for the tens. Each outcome can be produced immediately, and hence no interim results must be retained. However, when there is a carry, an extra operation is required. For example in the sum 38 + 24, the problem part of the units, 8 + 4, has 12 as the result. This result must be subdivided in the part of the tens (1) that must be retained (and carried) and the part of the units (2) that can be emitted immediately. Next, the part retained (the carry) can be used to perform the calculation on the tens: 1 (carry) + 3 + 2. It is transparent from this example that the carry is maintained in memory for a very brief period of time. For sure, another strategy can be followed, where the value of the carry is maintained in memory while the sum of the present column is calculated, and at the end of this calculation the maintained value is also added. In our example that would imply that the carry is maintained in memory while the addition of 3 + 2 is made and that afterwards the carry is added (5 + 1). The latter strategy obviously requires the maintenance of an interim result. We assume that, because it unnecessarily increases the memory load, this strategy is not used, so that we can focus on the simpler situation where the time during which the value of the carry must be maintained is rather limited, namely only during the response preparation interval for the answer part being emitted.

Manipulation of the number of carries thus increases the number of times this extra operation and transfer of the carry must be executed. Consequently, it is expected that with every carry the solution duration of the complete sum will be increased. With the increase in the number of operations performed, the probability of an error will also increase. Both these predictions are consistent with the data available in the literature (Ashcraft & Kirk, 2001; Fürst & Hitch, 2000; Seitz & Schumann-Hengsteler, 2002).

The value of the carry depends on how many numbers are being added. With two numbers, the value of the carry is either 0 (no carry) or 1. With three numbers, the maximum value of the carry is 2. Hence, manipulation of the value of the carries goes hand in hand with an increase in the number of column-wise additions. However, within the n numbers to be added, the value of the carry can vary from 0 (e.g., 2 + 3 + 3) to $n - 1$ (e.g., 8 + 7 + 9). Therefore, not the number of operations (which is the same in both examples), but the total amount of the outcome varies with the value of the carry. From the literature on simple mental arithmetic it is known that solution time and accuracy depend on the size of the outcome (the so-called problem-size effect; Ashcraft, 1992, 1995; Ashcraft & Battaglia, 1978; Butterworth, Zorzi, Girelli, & Jonckheere, 2001; Campbell, 1995; Geary, 1996; Groen & Parkman, 1972). Consequently, it may be predicted that the larger the value of the carry or carries, the slower the solution and the less accurate on average.

When both the number of carries and the value of the carries are varied orthogonally, the answer time will slow down because with more carries more operations have to be executed, and with increasing value of the carries, each of the partial outcomes will take more time. This results in a multiplicative combination of the two effects so that an interaction of these two manipulations is expected. As regards accuracy, however, with more carries there are more chances of committing an error and likewise with an increase in the value of the carries. The presence of more carries will not increase the probability of committing an error because of the value and vice versa, so that no interaction is expected.

In the present study, the contribution of two aspects of working memory is targeted. On the one hand, it is well known from studies with simple mental additions (De Rammelaere et al., 1999, 2001; De Rammelaere & Vandierendonck, 2001; Lemaire et al., 1996; Seitz & Schumann-Hengsteler, 2002) that an executive load slows down solution time and increases the proportion of errors. As explained above, the problems studied may be considered as a concatenation of simple arithmetic problems, and therefore it is expected that an executive load will affect both solution time and accuracy. Since an executive load will interfere with each calculation step done, this leads also to the prediction of an interaction with number of carries: Since every carry increases the number of steps and since each step may be affected by the executive load, it is expected that the load effects will be larger when there are more carries. As far as value of the carries is concerned, an interaction with an executive load may be expected if the effects of such a load are augmented with problem difficulty. The evidence in favour of a load by problem difficulty interaction is rather scanty: Some studies did not find such an interaction (e.g., De Rammelaere et al., 1999, 2001; De Rammelaere & Vandierendonck, 2001), while others did (e.g., Ashcraft & Kirk, 2001). For that reason, as higher values of the carry go hand in hand with larger outcomes, an interaction with load may be expected if the problem size is sufficiently variable.

Another aspect of working-memory load concerns phonological memory, which can be impaired by means of articulatory suppression. Thus far, the evidence indicates that in simple mental arithmetic this manipulation does not affect performance (see review by DeStefano & LeFevre, 2004). By consequence, an effect of articulatory suppression is only expected if the problem requires the participant to maintain interim information. As already pointed out, the amount of load on this aspect of working memory (the maintenance of the value of the carry for brief periods of time) is rather limited, so that a small amount of interference due to such a phonological load may expected. Since this load is present for short periods that do not overlap in time, the number of carries will not increase this amount of load, and hence no interaction with number of carries is expected. Similarly, there is no reason to assume that the amount of phonological load varies with the value of the carry to be maintained, and no interaction of value of the carry with a phonological working-memory load is expected.

A further characteristic of the problems considered here is that participants produce their answers incrementally online so that besides global solution times and accuracy also the solution times and accuracy of each step in the production of the answer can be subjected to an analysis. In general, it may be expected that each step in the production of the answer will take longer the more operations there are, and the larger the outcome of the operations will be. In particular, this means that a calculation that does not result in a carry to the next problem part will be faster than a calculation that includes such a carry. Similarly, when the problem part considered receives a value carried from the previous part, the number of operations is increased, and this will also result in a slower response. Furthermore, if the outcome of an answer part is larger, this will also slow down responding. In all these cases, the situations that lead to slower responding also increase the probability of an error, so that basically all the variations that slow down solution, on average, also result in decreased accuracy.

Pursuing this issue further, the number of operations performed on the units will on average be smaller than the number of operations performed on the tens. For the units, the *n* numbers have to be added, and occasionally the outcome will require a carry. The same may happen for the tens, but additionally a value may be carried from the units, so that on average more operations are required. If the numbers to be added are four-digit numbers, then no difference is expected between tens and hundreds, because on average the same number of operations will have to be executed. There will be a difference, however, between the last and last-but-one problem part (i.e., between hundreds and thousands in four-digit numbers and between tens and hundreds in three-digit numbers). This is because the sums that we are considering here are constructed in such a way that the last part itself never results in a carry, so that the outcome will always be rather small (smaller than 10) as only the digits have to be added together with the occasional carry from the previous part. In sum, it is expected that answers to the unit part will be faster and more accurate than the answers to the other parts and that the answer to the last part will be faster and more accurate than the answer to the problem parts in between.

Given the findings regarding executive load in simple arithmetic, it is expected that under executive load, each of the problem parts will be slower, and if the range of problem variation is large enough an interaction of executive load with outcome size of the problem parts may occur. With respect to a phonological load, the evidence suggests that such a load does not affect simple arithmetic, and consequently no effect of a phonological load is expected, and such a load is also not expected to interact with the size of the partial problems.

EXPERIMENT 1

In order to test all these predictions, in the first experiment, we orthogonally manipulated the number of carry operations (one or two) and the value of the carry (1 or 2) and combined this with four different memory loads. The experiment was designed to mirror as closely as possible the design used by Fürst and Hitch (2000). For that reason, the same three memory load conditions were included: namely, control (arithmetic only), phonological load (articulatory suppression), and executive load (Trails task). However, because the load imposed by the Trails task was considered to be very large and because this task also seems to load in an important way on phonological working memory, another executive load condition was included: namely, a condition with a continuous choice reaction time task with randomly spaced interstimulus intervals (CRT-R). In this task, a random sequence of low and high tones is presented at a random rate, and the participants are required to respond quickly by a corresponding keypress to each tone. Szmalec, Vandierendonck, and Kemps (2005) report a series of experiments showing that this task creates a moderate executive load while the effects of this load are dissociated from the effects of operations known to affect the slave systems in the model of Baddeley and Hitch (1974). The usefulness of the task has meanwhile been confirmed in several studies with mental arithmetic (e.g., Deschuyteneer & Vandierendonck, 2005a, 2005b; Deschuyteneer, Vandierendonck, & Muyllaert, in press; Imbo, Vandierendonck, & Vergauwe, in press).

Participants were required to add three 3-digit problems that resulted in a 3-digit answer by typing their answer in the order units–tens–hundreds. This way a fourth independent variable, problem component, was included.

Method

Participants and design

A total of 20 first-year psychology students (19 women and 1 man) of Ghent University (Belgium) participated for course requirements and credits. They were assigned to four counterbalancing conditions, which determined the order in which they performed the four (working-memory load) conditions of the experiment.

Materials

A set of 60 addition problems was constructed: 40 experimental stimuli and 20 fillers. Each problem consisted of three 3-digit numbers that summed to another 3-digit number. For the experimental stimuli, there were four categories: (a) one carry operation of value 1—for example, 175 + 311 + 307 = 793; (b) one carry operation of value 2—for example, 164 + 281 + 260 = 705; (c) two carry operations both of value 1—for example, 153 + 286 + 341 = 780; and (d) two carry operations both of value 2—for example, 145 + 187 + 378 = 710.

There were two types of filler item. The first category had no carry (number = 0, and value = 0). In the second category, problems had two carries, one with value 1 and one with value 2.

The digit 9 was excluded in all three problem digits in order to avoid ambiguous errors (see Fürst & Hitch, 2000). All categories were matched for problem size. Independent t tests revealed no differences between the problem sizes of all categories (all $p > .20$). For the categories (a) and (b), half of the problems had a carry from the units to the tens, and half had a carry from the tens to the hundreds. The problems were arranged in four blocks (i.e., the four memory load conditions), so that in each condition 15 problems (3 practice, 4 fillers, and 8 experimental problems) were presented. The different problem types were represented proportionally in each condition.

Procedure

We followed the procedure of Fürst and Hitch (2000) as closely as possible so that potential differences could not be explained by procedural discrepancies. All participants were tested individually in a quiet room. The same experimenter was present during the experimental session. Each problem was shown at the centre of a computer screen in column-wise Arabic notation. The problem remained visible until the participant responded.

The participants were told that on each trial, they would see three 3-digit numbers of which the correct sum was another 3-digit number. Participants were asked to type in the correct answer by first typing the units, then the tens, and finally the hundreds. In this way, strategy variability due to the use of different strategies was eliminated (e.g., Hitch, 1978). Participants saw the digits appear on the screen as they typed, and they were to complete their answer by pressing the enter-key. Response time was measured in milliseconds as the time between the start of problem presentation and the completion of the answer (enter-key). The intertrial interval was 1,000 ms.

A dual-task selective interference methodology was used with one single-task control condition and three conditions in which participants performed the arithmetic task concurrently with a secondary task taxing a particular component of working memory. In the first dual-task condition, participants solved the arithmetical problems while continuously saying “de” (Dutch for “the”), at a rate of about 2 words per second (articulatory suppression). This task was meant to interfere with the rehearsal mechanism of the phonological loop. The second dual-task condition combined the arithmetic task with the Trails task. Participants were given a random letter and a random day of the week (e.g., “D-Friday”). The participants were requested to continue this series by alternating between the letters and the days of the week. When the end of the series was reached, the sequence had to be continued from the beginning (from Sunday to Monday and from Z to A). Baddeley (1996) has shown that this task interferes with executive functioning because switching between familiar streams probably requires the inhibition of prepotent responses. In the third dual-task condition, participants performed the arithmetic task concurrently with a continuous two-choice reaction time task with random spacing of the interstimulus interval (CRT-R). The interstimulus intervals were either 900 or 1,500 ms, and the stimuli were randomly selected from two tones: a low tone (262 Hz) and a high tone (524 Hz). The participants had to say “hoog” (“high”) when they heard a high tone and “laag” (“low”) when a low tone was presented. The duration of each tone was 200 ms. In a series of experiments, Szmalec et al. (2005) have shown that this task taxes executive functioning, while the load on the subordinate working-memory systems is negligible.

In order to familiarize the participants with the apparatus and the procedure, the experiment started with a few practice problems. After the explanation of the secondary task, the execution of the primary task in combination with the secondary task was practised too. After these practice problems, the blocks with the experimental items were presented. The four load conditions were presented in the order determined by a randomized Latin square, and within each block the problems were presented in a random order. In the dual-task conditions, the participants first started execution of the secondary task, and after 5 seconds the first arithmetical problem appeared, and the participants had to solve this problem while continuing with the secondary task. In the control condition, the first arithmetical problem was presented as soon as the participant was ready.

Performance on all secondary tasks was measured. The spoken responses of the participants in the articulatory suppression condition and in the Trails condition were tape-recorded and were analysed afterwards. For the CRT-R task, the experimenter checked online whether the responses of the participants were correct. The participants also performed the secondary tasks alone for 2 minutes ("single secondary task control condition"). Performance in these conditions was also measured.

Results

All repeated measures analyses in this and in the next experiment were performed by means of a multivariate analysis based on the multivariate linear model. The analyses of the primary task data—reaction time (RT) and accuracy—were further refined by means of regression analyses, and analyses of secondary task performance are reported as a check on possible dual-task trade-offs. For the statistical tests an α-level of .05 is assumed, unless otherwise mentioned.

Solution time

The complete factorial design was a 4 (load: control, articulatory suppression, CRT-R and Trails) $\times$ 2 (number: 1 or 2 carries) $\times$ 2 (value: 1 or 2) $\times$ 3 (components: unit, ten, hundred response component) with repeated measures on all the effects. Table 1 presents the average solution time as a function of load, number, value, and component. Only the correctly solved sums were included in these analyses.[1]

The first issue addressed in this analysis concerns only the effects of number and value. In order to ascertain that the effects were not due to the dual-task conditions, two analyses were carried out: (a) restricted to the control condition and (b) averaged over all four conditions. Including only the control condition, the effects of number and value as well as their interaction were significant: respectively, $F(1, 19) = 18.35$, $F(1, 19) = 26.65$, and $F(1, 19) = 5.55$. In the analysis of the complete data-set the pattern of findings was similar. Problem solving was slower with two carries (5.29 s) than with one carry (4.25 s): $F(1, 19) = 56.07$. Similarly, problems with a carry of value 2 were solved slower (5.14 s) than problems with a carry of value 1 (4.40 s): $F(1, 19) = 10.06$. The interaction of both effects was significant as well, $F(1, 19) = 11.68$, in such a way that problems with two carries of value 2 (5.99 s) were solved slower than the others (4.21–4.59 s).

The second part of the analysis showed that load also affected arithmetic performance, as expected, $F(3, 17) = 44.95$. Reaction times per component were slower under an executive load (CRT and Trails; 6.34 s) than under articulatory suppression (3.34 s), $F(1, 19) = 140.73$, and in the latter condition performance was slower than that in the control condition (3.07 s), $F(1, 19) = 7.16$. Consequently, reaction time in the control condition was also faster than that in the conditions with an executive load, $F(1, 19) = 143.01$. For completeness, performance also differed between the CRT-R (5.05 s) and the Trails task (7.62 s), $F(1, 19) = 19.48$.

[1] The inclusion of only the solution times of the correct sums resulted in a number of empty cells (i.e., where no problem of that type was solved correctly). We replaced these empty cells (126 or 13% in Experiment 1; 264 or 12% in Experiment 2) with the mean of the Load $\times$ Number $\times$ Value $\times$ Component cell (see, e.g., Roth, 1994).

Table 1. *Mean solution times*[a] *and standard errors as a function of load, number of carry operations, value of the carry, and response components in Experiment 1*

		Number of carry operations			
		1		*2*	
Load	*Component*	*Value 1*	*Value 2*	*Value 1*	*Value 2*
Control	Unit	2.86 (0.30)	2.85 (0.21)	2.82 (0.29)	4.64 (0.43)
	Ten	2.74 (0.28)	3.34 (0.28)	2.99 (0.22)	3.89 (0.28)
	Hundred	2.14 (0.15)	2.77 (0.41)	2.93 (0.27)	2.89 (0.20)
Phonological	Unit	2.57 (0.27)	3.13 (0.24)	3.04 (0.25)	4.61 (0.45)
	Ten	2.45 (0.23)	4.20 (0.39)	3.40 (0.31)	4.53 (0.41)
	Hundred	2.65 (0.33)	2.71 (0.31)	2.80 (0.18)	3.92 (0.35)
Executive (CRT-R)	Unit	4.39 (0.54)	4.33 (0.84)	4.53 (0.57)	6.33 (0.78)
	Ten	4.17 (1.11)	4.17 (0.45)	6.83 (1.62)	6.85 (1.03)
	Hundred	4.32 (0.62)	3.38 (0.41)	4.86 (0.83)	6.45 (1.08)
Executive (Trails)	Unit	8.09 (1.31)	6.85 (0.52)	6.96 (0.96)	8.46 (0.81)
	Ten	7.14 (1.28)	7.02 (0.96)	6.94 (0.76)	10.42 (1.17)
	Hundred	6.97 (1.12)	6.74 (0.96)	7.02 (1.13)	8.84 (1.55)

Note: Standard errors in parentheses.
[a]In s.

Next, we addressed the interactions of load with number and value. Although the effect of number was larger in the conditions with a memory load (a difference of 1.19 s) than in the control condition (difference 0.58 s), the overall interaction was only marginally significant, $F(3, 17) = 2.96$, $p = .06$. The interaction between load and value turned out not be significant, $F(3, 17) = 1.54$, $p = .24$.

A further decomposition of the interaction of load and number was performed to clarify the roles of the phonological and the executive loads. This analysis showed that the interaction of number with the contrast between control and articulatory suppression was not significant, $F < 1$, whereas number interacted with the contrast between control and executive loads, $F(1, 19) = 6.83$. The contrast between the two executive conditions did not interact with number, $F(1, 19) = 1.49$, $p = .24$. It thus seems that the marginal overall interaction is completely due to the presence of an executive load.

Finally, the effects related to the problem components are reported. The main effect of components fell short of significance, $F(2, 18) = 2.97$, $p = .08$. Neither of the interactions between components and the other effects attained significance (smallest $p = .14$ for the quadruple interaction of load, number, value, and components).

Thus far, the results with respect to the solution time data seem to indicate that solution time depends on problem characteristics such as number of carries and value of the carry, on the one hand, and working-memory limitations, on the other hand. Working-memory load, however, seems to interact only with the number of carries. Because the problems used were randomly selected from a number of predefined categories, additional regression analyses were performed as suggested in Method 3 of Lorch and Myers (1990). It is well known that solution times depend on problem size. Since the outcomes of the problems were balanced over the number by value conditions the variable of overall problem size (outcome) does not reflect the problem difficulties due to the carries. A better variable is obtained by taking the column-wise sums. With the addition 172 + 235 + 284, the outcomes of the units, tens, and hundreds are, respectively, 11, 19 (18 + carry), and 6 (5 + carry). The sum of these outcomes is a better indicator of problem difficulty, as this sum will be greater the more carries are present, the larger the value of the carries, and the larger the column-wise sums. This measure of "problem size" was entered as a predictor into regression analyses per participant. Because this measure of problem size correlates .96 with the interaction of number and value, an inclusion of all effects of the model in the regression analyses would lead to singularities. Moreover, in the present experiment, the number of useful observations per participant was rather low. Therefore, it was decided to include only load and value as the other predictors. Because load is a categorical variable it was coded by means of three dummy variables. The reaction times averaged over the three components were regressed on these predictors per participant, and the regression coefficients were entered in a multivariate analysis of variance. This analysis showed that the regression coefficients associated with these predictors were not significant, $F < 1$ for problem size, $F(3, 17) = 1.42$, $p = .27$ for load, and $F(1, 19) = 2.37$, $p = .14$ for value. A similar analysis with load, number, and size as predictors yielded similar results: respectively, $F(3, 17) = 2.58$, $p = .09$, $F < 1$, and $F < 1$.

Accuracy

The same factorial design was used for the analysis of the accuracy data. Mean accuracy is shown in Table 2 as a function of load, number, value, and component. The results are reported in the same order as that for the reaction times.

The effects of number and value and their interaction are first considered within the control condition only. Both main effects were significant, $F(1, 19) = 5.59$ for number, and $t(19) = 1.81$, one-tailed, for value. Their interaction was not significant, $F < 1$. In the overall analysis, the pattern was completely similar, but the effects were slightly stronger: Both main effects were significant, $F(1, 19) = 7.44$ for number, and $F(1, 19) = 10.83$ for value, but their interaction was not, $F < 1$. Accuracy was higher in problems with one (.90) than in problems with two carries (.85). Accuracy was also higher when the value of the carries was one (.91) than when it was two (.84).

The second part of the analysis again focused on the main effect of memory load. As expected, this effect was reliable, $F(3, 17) = 11.10$.

Table 2. *Mean proportions of correctly solved sums and standard errors as a function of load, number, value, and component in Experiment 1*

		Number of carry operations			
		1		*2*	
Load	*Component*	*Value 1*	*Value 2*	*Value 1*	*Value 2*
Control	Unit	1.00 (.00)	.93 (.04)	.95 (.03)	.90 (.04)
	Ten	.98 (.02)	.93 (.04)	.90 (.04)	.88 (.05)
	Hundred	.98 (.02)	.93 (.04)	.93 (.04)	.93 (.04)
Phonological	Unit	.95 (.03)	.95 (.03)	.95 (.03)	1.00 (.00)
	Ten	1.00 (.00)	.93 (.06)	.85 (.06)	.88 (.06)
	Hundred	.93 (.04)	.90 (.06)	.93 (.05)	.88 (.05)
Executive (CRT-R)	Unit	1.00 (.00)	.90 (.04)	.95 (.03)	.90 (.04)
	Ten	.93 (.04)	.75 (.09)	.88 (.05)	.75 (.08)
	Hundred	.93 (.04)	.75 (.07)	.88 (.05)	.78 (.05)
Executive (Trails)	Unit	.95 (.03)	.93 (.05)	.95 (.03)	.98 (.02)
	Ten	.80 (.07)	.73 (.07)	.65 (.07)	.48 (.08)
	Hundred	.83 (.07)	.70 (.07)	.78 (.07)	.55 (.09)

Note: Standard errors in parentheses.

Planned comparisons on this effect showed that the control (.93) and the articulatory suppression condition (.93) did not differ from each other, $F < 1$, but the contrast between articulatory suppression and executive load (.82) was significant, $F(1, 19) = 27.20$. Compared to the control condition, the effect of an executive load was also significant, $F(1, 19) = 28.29$. Again the effect of the Trails task was larger (.78) than that of the CRT-R task (.86), $F(1, 19) = 9.32$.

Next we considered the interactions of load with number and value. Only the interaction of load by value was reliable, $F(3, 17) = 5.14$, while the interaction of Load × Number and the triple interaction of load, number, and value were not, both $F < 1$. Decomposition of the interaction of load and value showed that the contrast between control and articulatory suppression did not attain significance, $F(1, 19) = 1.15$, while the contrast between control and executive load did, $F(1, 19) = 4.26$.

The last aspect concerns the response components. The main effect of component was significant, $F(2, 18) = 31.29$. Accuracy was largest for the units (.95), smallest for the tens (.83) and in between for the hundreds (.85). The contrast between unit and the other components was significant, $F(1, 19) = 51.75$, but the difference between the ten and hundred response was not, $F(1, 19) = 1.67$, $p = .21$. The effect of component interacted with number, $F(2, 18) = 7.96$. This was essentially based on the interaction of the contrast between units and the other components with number, $F(1, 19) = 10.31$. The interaction of component with value failed to attain significance, $F(2, 18) = 3.11$, $p = .07$, but here also the contrast between units and the other components interacted with value, $F(1, 19) = 5.25$. Component interacted also with load, $F(6, 14) = 8.28$. The contrast between units on the one hand and tens and hundreds on the other interacted with the presence of an executive load (control vs. executive load), $F(1, 19) = 18.54$, but not with the contrast between control and phonological load, $F(1, 19) = 1.74$, $p = .20$. This seems to indicate that the difference in accuracy between units and the other answer components is augmented under an executive load: a difference of .02 in the control condition versus a difference of .05 in the condition with a phonological load and a difference of .19 in the conditions with an executive load. Moreover, the interaction with the contrast between the two executive load conditions was also significant, $F(1, 19) = 24.02$, with differences of, respectively, .11 and .26 in the CRT-R and Trails conditions. If it can be assumed that the executive load is larger in the Trails than in the CRT-R task, then this would indicate that the stronger the load, the larger the difference in accuracy between the response components.

The results of the analysis of the accuracy data suggest that accuracy to some extent depends on the value of the carries and on the available memory capacity. Therefore, regression analyses were performed as suggested in Method 3 by Lorch and Myers (1990). Problem size was measured in the same way as for the analysis of the solution time data and was entered as a predictor in the regression analyses together with load and value. As before, load was coded by means of orthogonal 3 dummies. The participant's average accuracy over the three response components was regressed on these predictors, and the regression coefficients were entered in a multivariate analysis of variance to estimate the contribution of these predictors. Size and value were not significant, $F(1, 19) = 3.53$, $p = .08$, and $F(1, 19) = 2.50$, $p = .13$, respectively, but load was, $F(3, 17) = 11.05$. A similar analysis with load, number, and size as predictors showed significant effects of load, $F(3, 17) = 11.21$, and size, $F(1, 19) = 5.46$, but not of number, $F < 1$.

Secondary task performance

In the articulatory suppression condition, participants did not slow down their speech rate during calculation as compared to the secondary task-only control condition: respectively, 107 and 103 words per minute, $t(19) = 1.21$, $p > .20$. Performance on the Trails task was measured by the number of responses produced per minute and by the proportion of errors. Response production was lower in the dual-task (27) than in the single-task (52) condition, $t(19) = 12.56$.

Participants also committed proportionally more errors (3.3%) in the dual-task than in the single-task condition (1.6%), $t(19) = 2.85$. For the CRT-R task, the same pattern of results was observed: Participants committed more errors during calculation (14%) than in the CRT-R-only condition (4%), $t(19) = 4.18$. All these findings show that not only performance of the primary task, but also performance of the secondary tasks was impaired in the dual-task conditions.

Discussion

Both the number of carries and the value of the carries affected arithmetic performance such that more carries and higher values of the carry resulted in poorer performance. As predicted, the two factors interacted in the solution times, but not in accuracy. Problems with two carries of value 2 were solved much slower than the other problems. In fact this pattern of results might be due to the presence of memory load conditions. Therefore, we also investigated these effects restricted to the control condition. Basically, the same pattern was found.

The presence of memory loads impaired performance, and this effect was larger under an executive load than under articulatory suppression. Within the executive load, it can be seen that the Trails task had a more devastating effect than the CRT-R task. Load interacted with number and value, although the pattern was not completely as expected. In the solution times, load interacted with number but not value. Response times were longer with more carries in all the load conditions, but the difference was larger in the conditions with an executive load, while the difference was smaller and approximately the same in the control and the articulatory suppression condition. The interaction of memory load and value was significant in the accuracy data only, and it appeared that the effect of value was very strong in the executive load conditions while absent in the other conditions (articulatory suppression and control).

The responses for each of the answer components were about equally fast, and these partial solution times were not moderated by any of the effects. However, accuracy was very different over the answer components. As expected, the accuracy of the units was greater than that of the other components. This difference was augmented by an executive load and by the presence of more carries in the problem.

Additional regression analyses were performed. These analyses showed that solution times depended on memory load when problem size (the sum of all elementary outcomes), was also included as a predictor. The number of data points available was too small, however, for a powerful test of these effects.

In sum, the findings of the first experiment confirm the expectation that besides number of carries, value of the carry also plays an important role in arithmetic performance. While the present experiment confirmed the interaction of memory load with number of carries only for solution times, it was found that the effect of value of the carries seems to be augmented by the presence of an executive load and not by a phonological load, but only in the accuracy data. Taken together, this experiment confirms an important role for the value of the carries but leaves some room for doubts about the mechanism underlying this effect. In particular, the present findings do not clarify whether the effect of value of the carries stands on its own or rather follows from the interaction with number of carries. Furthermore, more data would be useful to clarify how number and value of the carry interact with memory load. These issues were further pursued in Experiment 2.

EXPERIMENT 2

In the second experiment, the value of the carry was varied in the range 1 to 3. To achieve this, it was necessary to increase the number of elements to be added from 3 to 4. In order not to restrict the problem composition too much, it was also decided to increase the number of digits in each number from 3 to 4. This way, it was possible also to extend the range of the number of carries, such that problems could be presented with 1 to 3 carries. Finally, because the load due to the Trails task was already at its limits in Experiment 1, it was decided not to include the Trails condition in Experiment 2 in order to avoid a complete break down of calculation. Only the CRT-R task was used to create an executive load condition.

Method

Participants and design

A total of 20 first-year psychology students of Ghent University (Belgium) participated for course requirements and credit: 3 men and 17 women. Their mean age was 18.3 years. None of them had participated in the first experiment.

Materials and procedure

A set of 90 addition problems was constructed: 72 experimental stimuli and 18 fillers. The experimental items consisted of four 4-digit numbers that summed to obtain another 4-digit number. The number of carry operations was one, two, or three; the value of the carry was 1, 2, or 3. This resulted in nine categories:

(a) one carry operation of value 1—for example, 2536 + 1621 + 2320 + 1121 = 7598;
(b) one carry operation of value 2—for example, 5112 + 1225 + 2418 + 1207 = 9962;
(c) one carry operation of value 3—for example, 1831 + 1804 + 2721 + 1812 = 8168;
(d) two carry operations of value 1—for example, 1162 + 2872 + 2101 + 2321 = 8456;
(e) two carry operations of value 2—for example, 1138 + 4086 + 3346 + 1173 = 9743;
(f) two carry operations of value 3—for example, 4088 + 1177 + 1477 + 1068 = 7810;
(g) three carry operations of value 1—for example, 1623 + 2526 + 4172 + 1143 = 9464;
(h) three carry operations of value 2—for example, 1457 + 2306 + 1584 + 1854 = 7201;
(i) three carry operations of value 3—for example, 2588 + 1788 + 1778 + 1878 = 8032.

For each category, eight problems were constructed, distributed over sums in the seven thousands, in the eight thousands, and in the nine thousands. This way, all types were matched for problem size. Independent t tests revealed no differences between the problem sizes of all types. As in Experiment 1, we controlled for the place of the carry operation. The digit 9 was excluded in all four problem digits to avoid ambiguous errors (see Fürst & Hitch, 2000). The 18 filler items consisted of problems with two or three carry operations, with mixed values for the carries (e.g., one carry with a value of 1, one carry with a value of 2, and one carry with a value of 3), and were not included in the analyses.

Apparatus and procedure were identical to those of the first experiment, with only two differences. As already mentioned, we skipped the Trails condition. Thus, only three conditions, of which the order was counterbalanced, were included: control, articulatory suppression, and CRT-R task. The second difference is that the participants did not have to press the enter key any more; the problems succeeded each other automatically.

Results

The same data-analytic strategy was applied as in Experiment 1. Unless otherwise stated, results are significant at $\alpha = .05$.

Solution time

The solution times were analysed on the basis of a 3 (load: control, articulatory suppression, CRT-R) $\times$ 3 (number: 1, 2, or 3 carries) $\times$ 3 (value: 1, 2, or 3) $\times$ 4 (component: unit, ten, hundred, thousand response) factorial design with repeated measures on all effects. The average solution times as a function of these effects are displayed in Table 3.

The same scheme for presentation of the results is followed as in Experiment 1. First, confined to

Table 3. *Mean solution times*[a] *and standard errors as a function of load, number of carry operations, value of the carries, and response components in Experiment 2*

			Number of carry operations								
			1			*2*			*3*		
Condition	*Component*		*Value 1*	*Value 2*	*Value 3*	*Value 1*	*Value 2*	*Value 3*	*Value 1*	*Value 2*	*Value 3*
Control	Unit	Mean	1.74	1.95	2.07	2.07	3.11	2.13	1.85	3.05	2.8
		SE	0.22	0.24	0.29	0.31	0.48	0.26	0.2	0.53	0.36
	Ten	Mean	1.92	2.22	1.94	2.47	2.93	2.98	2.04	3.23	3.85
		SE	0.16	0.32	0.3	0.37	0.4	0.44	0.19	0.32	0.45
	Hundred	Mean	1.89	2.71	1.88	2.08	3.48	2.77	2.49	2.99	5.27
		SE	0.17	0.36	0.26	0.19	0.46	0.37	0.28	0.34	0.49
	Thousand	Mean	1.77	2.02	1.8	1.86	2	2.11	1.95	2	2.1
		SE	0.19	0.23	0.19	0.15	0.21	0.23	0.25	0.17	0.15
Articulatory	Unit	Mean	2.88	2.27	2.05	1.92	3.46	2.8	2.03	3.24	2.91
suppression	Ten	*SE*	0.63	0.42	0.21	0.16	0.42	0.54	0.17	0.51	0.3
		Mean	2.57	3.11	1.87	2.08	3.87	3.1	2.8	3.25	4.94
		SE	0.45	0.46	0.19	0.3	0.61	0.45	0.36	0.35	0.51
	Hundred	Mean	2.46	2.72	2.01	2.78	3.67	4.07	3.25	3.55	5.43
		SE	0.24	0.37	0.36	0.42	0.59	0.84	0.37	0.76	0.54
	Thousand	Mean	2.45	1.92	2.23	2.16	3.96	2.38	2.31	2.27	2.32
		SE	0.35	0.26	0.18	0.34	1.11	0.38	0.18	0.23	0.24
CRT-R:	Unit	Mean	2.79	2.38	3.78	3.55	5.1	3.67	4.38	4.48	2.28
Executive load		*SE*	0.42	0.36	0.8	0.41	0.76	0.57	0.71	0.62	0.24
	Ten	Mean	3.4	4.79	3.65	4.05	4.46	4.76	3.16	4.74	6.51
		SE	0.63	1.39	0.64	0.82	0.7	0.84	0.39	0.77	0.84
	Hundred	Mean	3.04	3.34	2.93	4.37	9.54	7.51	3.92	5.19	8.52
		SE	0.36	0.43	0.4	0.62	2.62	1.68	0.6	0.88	1.3
	Thousand	Mean	3.75	4.06	2.95	3.66	3.05	3.83	4.34	3.71	3.46
		SE	1.26	0.72	0.39	0.47	0.31	0.49	0.98	0.37	0.63

[a]In s.

the control condition, the effects of number and value and their interaction were all significant: respectively, $F(2, 18) = 7.83$, $F(2, 18) = 8.45$, and $F(4, 16) = 3.53$. The same pattern occurred in the overall analysis, with, respectively, $F(2, 18) = 19.88$, $F(2, 18) = 12.08$, and $F(4, 16) = 6.46$. The main effects were further decomposed by means of contrasts. For number, solution time was shorter for one carry (2.59 s) than for multiple carries (3.56 s), $F(1, 19) = 40.81$, but the difference between problems with two (3.44 s) and three carries (3.52 s) was not significant, $F < 1$. Similarly, problems with carries of value 1 (2.73 s) were answered faster than problems with larger values (3.41 s), $F(1, 19) = 22.21$, and there was no difference when the value was 2 (3.44 s) or 3 (3.38 s), $F < 1$. The interaction of number and value is shown in Figure 1. The figure shows that when both number and value are larger than 1, the problems are solved slower. This is confirmed in the interaction of the contrast 1 versus more on both effects, $F(1, 19) = 7.97$. Furthermore, value 2 versus 3 interacted with number 2 versus 3, $F(1, 19) = 4.94$. This is seen in the figure in the cross-over of the lines for values 2 and 3.

The second aspect of our data analysis concerns the effect of load, $F(2, 18) = 15.74$. Responses were faster in the control (2.43 s) than in the articulatory suppression condition (2.86 s),

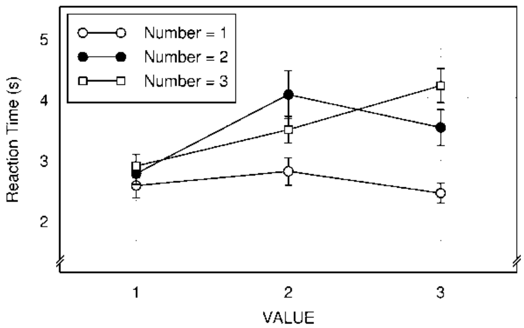

Figure 1. *Interaction of number by value in the reaction time data of Experiment 2.*

$F(1, 19) = 6.41$, and than in the CRT-R condition (4.25 s), $F(1, 19) = 33.22$. The difference between the articulatory suppression and the CRT-R conditions was also significant, $F(1, 19) = 24.30$. Load did not interact with either number or value, all $F < 1$.

In contrast to Experiment 1, the effect of component was significant, $F(3, 17) = 32.78$. Responses for the units were faster (2.84 s) than the other responses (3.30 s), $F(1, 19) = 12.89$. Responses for the tens (3.36 s) were not different from the other responses, $F < 1$, but responses to the hundreds were slower (3.85 s) than the responses to the thousands (2.68 s), $F(1, 19) = 88.41$. Component interacted with number, with value, and with their interaction: respectively, $F(6, 14) = 12.92$, $F(6, 14) = 20.24$, and $F(12, 8) = 3.63$. Component did not interact with load in any way. Further exploration of the interaction with number showed that number interacted with the contrast between hundreds and thousands, $F(2, 18) = 21.43$. The difference in RT for hundreds versus thousands was only .01 s if number was 1, but it amounted to 1.74 s when number of carries was more than 1. The interaction of component with value was decomposed also on the basis of the three orthogonal contrasts on component. The unit RT differed more from the other RTs when the value was larger, $F(2, 18) = 7.72$. Similarly, the RT difference between hundred and thousand was larger with larger carry values, $F(2, 18) = 15.85$.

Similar regression analyses were performed to those in Experiment 1, according to the methodology described in Method 3 of Lorch and Myers (1990). Problem size was defined in the same way as the sum of the outcomes of the four columns. Because more data points per participant were available than in the previous experiment, it was possible to test the effect of problem size together with the other predictors based on the factorial design. However, the predictor Number × Value was not included because it correlates .98 with problem size. With problem size, load, number, value, Load × Number, and Load × Value as predictors, problem size, load, and value were significant, $F(1, 19) = 23.73$, $F(2, 18) = 15.37$, and $F(1, 19) = 4.86$, respectively, as was the interaction of load by value, $F(2, 18) = 4.24$. The effects of number and load by number were not significant, both $F < 1$.

Accuracy

The accuracy data were analysed by means of the same design. The average proportions correct and their standard errors as a function of load, number, value, and component are shown in Table 4.

Number of carries and value of the carry had again reliable effects, $F(2, 18) = 11.19$, and $F(2, 18) = 6.56$, respectively. With two or more carries accuracy was smaller (.87) than when only one carry was present (.92), $F(1, 19) = 17.90$, and the difference between two (.89) and three carries (.85) was also reliable, $F(1, 19) = 9.79$. Problems with carries of value 1 were more often correct (.92) than problems with carries of value 2 or more (.87), $F(1, 19) = 11.70$, and the difference between problems with carries of value 2 (0.89) and 3 (0.85) was also reliable, $F(1, 19) = 6.67$. These two main effects did not interact, $F(4, 16) = 1.93$, $p = .16$. Confined to the control condition, neither of the effects of number and value, nor their interaction attained significance. Nevertheless, for the effect of number, the pattern of findings was similar to that in the complete experiment as shown by the significant linear trend on number of carries, $F(1, 19) = 5.16$. For value,

Table 4. *Mean proportions of correct responses and standard errors as a function of load, number of carry operations, value of the carry, and response component in Experiment 2*

			Number of carry operations								
			1			2			3		
Condition	Component		Value 1	Value 2	Value 3	Value 1	Value 2	Value 3	Value 1	Value 2	Value 3
Control	Unit	Mean	.98	.95	1.00	.98	.93	.97	.98	.92	.93
		SE	.02	.03	0.00	.02	.04	.02	.02	.04	.04
	Ten	Mean	.93	.93	.98	.89	.95	.88	.91	.92	.83
		SE	.04	.03	.02	.04	.03	.05	.05	.04	.05
	Hundred	Mean	.93	.96	.96	.93	.92	.9	.96	.9	.84
		SE	.04	.03	.03	.04	.04	.05	.03	.04	.06
	Thousand	Mean	.96	.93	.94	.94	.98	.98	.95	.9	.96
		SE	.03	.04	.04	.03	.02	.02	.03	.04	.03
Articulatory suppression	Unit	Mean	.98	.96	.98	.96	.95	.96	.95	.88	.88
		SE	.02	.03	.02	.03	.03	.03	.03	.05	.06
	Ten	Mean	.9	.87	.86	.93	.88	.78	.83	.91	.67
		SE	.04	.05	.05	.04	.05	.05	.07	.04	.07
	Hundred	Mean	.9	.93	.83	.93	.87	.81	.82	.83	.63
		SE	.05	.04	.06	.04	.05	.06	.06	.05	.07
	Thousand	Mean	.88	.93	.93	.93	.86	.95	.88	.93	.68
		SE	.04	.03	.03	.04	.06	.03	.05	.04	.07
CRT-R: Executive load	Unit	Mean	1.00	.96	.88	.94	.9	.94	.93	.95	.89
		SE	0.00	.03	.06	.03	.04	.03	.05	.03	.04
	Ten	Mean	.9	.93	.85	.87	.77	.78	.86	.72	.65
		SE	.04	.04	.05	.05	.06	.07	.06	.08	.06
	Hundred	Mean	.93	.91	.83	.85	.84	.55	.86	.69	.57
		SE	.04	.04	.05	.05	.05	.08	.04	.08	.07
	Thousand	Mean	.93	.92	.83	.79	.81	.84	.95	.66	.86
		SE	.03	.04	.05	.06	.05	.05	.04	.09	.06

the averages showed the same trend, but this was not significant, $F < 1$.

Overall, load affected accuracy, $F(2, 18) = 16.90$. Accuracy was higher in the control condition (.94) than in the articulatory suppression condition (.88), $F(1, 19) = 18.22$, and than in the CRT-R condition (.84), $F(1, 19) = 32.00$. The two dual-task conditions also differed from each other, $F(1, 19) = 5.14$. Furthermore, load interacted with number and with value, respectively, $F(4, 16) = 4.41$, and $F(4, 16) = 3.04$. The interaction of load by number is displayed in Panel A of Figure 2. The figure clearly shows that both load conditions impair accuracy but only when multiple carries are present. This was tested by taking the interaction of control versus load by one versus multiple carries, $F(1, 19) = 8.00$. In order to disentangle the effects of phonological and executive loads, the interaction of the contrast of control versus articulatory suppression by number and the contrast of control versus CRT-R were tested. The latter was significant, $F(2, 18) = 5.40$, but the former was not $F(2, 18) = 2.23$, $p = .14$.

Panel B of Figure 2 shows the interaction of load by value. This interaction was studied in a similar way to the previous one. The contrast of control versus load interacted with the contrast of value 1 versus larger, $F(1, 19) = 7.30$. Value did interact with the contrast of control versus CRT-R, $F(2, 18) = 5.04$, but did not interact with the contrast of control versus articulatory suppression, $F(2, 18) = 2.25$, $p = .14$.

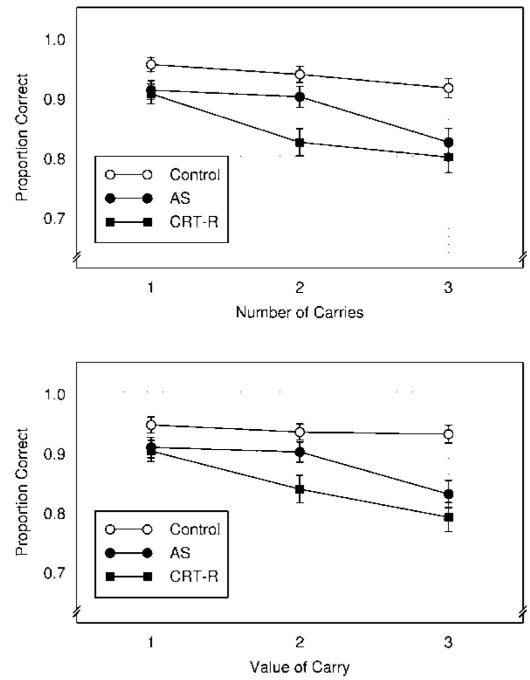

Figure 2. *Interactions of load in the accuracy data of Experiment 2. Panel A: Load by number. Panel B: Load by value. The labels control, AS, and CRT-R refer to, respectively, the control, articulatory suppression, and CRT-R conditions.*

The main effect of component was also significant, $F(3, 17) = 40.45$. Responses for the unit sum (.95) were more often correct than the other responses (.87), $F(1, 19) = 95.43$. Responses for tens (.86) and hundreds (.85) were equally accurate, $F(1, 19) = 1.01$, $p = .33$, and responses for thousands (.89) were more accurate than those of hundreds, $F(1, 19) = 26.30$. Component interacted marginally with value, $F(6, 14) = 2.72$, $p = .06$. Further decomposition of the latter interaction showed that the contrast of thousands versus hundreds interacted with value, $F(2, 18) = 7.94$.

Finally, regression analyses conform to Method 3 of Lorch and Myers (1990); with problem size as defined before, load, number, value, load by number, and load by value as predictors revealed significant effects of size, Load × Number, and Load × Value: respectively, $F(1, 19) = 5.48$, $F(2, 18) = 5.31$, and $F(2, 18) = 4.67$. The other predictors failed to attain significance, largest $F(2, 18) = 2.29$, $p = .13$ for load.

Error analysis

Because a sufficient number of errors were committed spread over the different conditions in the experiment, it was possible to perform an analysis of the errors to clarify the role of carrying. Error rate was largest in the tens and the hundreds (14% and 15%, respectively); it was lower in the thousands (11%) because of the more restricted possibilities (all sums were in the seven, eight, and nine thousands), and it was lowest in the units (6%).

We tested whether errors were due to a mistake in the carry procedure. Errors were indeed more frequent when carrying was needed: As well for the units, the tens, the hundreds, and the thousands, 66% of the errors occurred when carry operations were needed. In most of the cases, erroneous tens, hundreds, and thousands were lower than the correct tens, hundreds, and thousands, which suggests forgetting to carry or carrying too small a number. In this respect, carrying a number too small by 1 (a -1 error) was highly frequent, independent of the value that had to be carried (28% for value 1, 28% for value 2, and 44% for value 3). The -2 errors occurred especially in sums where a 2 or a 3 had to be carried (33% and 56%, respectively). The -3 errors almost only occurred when a 3 had to be carried (76%).

In summary, these analyses show that very often participants forgot to execute the carry operation and that when they did not forget, the errors were in the direction of carrying too small a number.

Secondary task performance

In the articulatory suppression condition, participants did not slow down their rate of saying "the" while calculating as compared to a single secondary task control condition: respectively, 90.6 and 85.3 words per minute, $t(19) = 1.00$. However, performance on the CRT-R task declined under dual-task conditions, $t(19) = 6.44$, with 97.4 correct responses when the task was performed in isolation, compared to 71.5 correct responses when performed in combination with the mental arithmetic task. The pattern of these results is quite similar to that of Experiment 1, and again under an executive load both the primary and the secondary task are impaired.

Discussion

The findings of this experiment can be summarized as follows. First, the pattern of results for number of carries, value of the carries, and their interaction replicated that of Experiment 1 and is consistent with the predictions formulated in the Introduction. More specifically, problems with one carry were more correct and faster than problems with multiple carries. Within the latter, problems with two carries were more correct but not faster than problems with three carries. In a similar pattern, problems with carries of value 1 were more correct and faster than problems with higher values of the carry, and problems with carries of value 2 were more correct but not faster than problems with carries of value 3. The interaction of these two factors boils down to slower performance when there are multiple carries of higher value. Or stated otherwise, if the number of carries is 1, the value does not matter much, and when the value of the carries is 1, the number of carries does not matter much (Figure 1). Taken all together, this subset of the findings shows that besides number of carries, the value is also an important determinant of calculation performance. At the same time, the effect of both factors must be qualified, because the effect of number becomes important only with higher values of the carry, and the effect of value of the carry is similarly amplified by the number of carries, at least when solution time is considered.

A second subset of findings concerns memory load and its interactions with number and value of carries. Calculation performance was impaired by concurrent articulatory suppression and was even more impaired by an executive load, which consisted of concurrently and continuously performing the CRT-R task. In the accuracy data, but not in the solution times, memory load interacted with number of carries and with the value of the carries. Figure 2 shows that these two interactions are very similar to each other. Indeed, for the interaction of memory load by number, the accuracy data show that performance is worst when there is a memory load combined with multiple carries. This figure also shows that under articulatory suppression the effect is smaller than that under executive load, which was confirmed by the absence of a significant effect of articulatory load in the data analysis. Similarly, proportion correct was adversely affected by load when the value of the carries was larger than 1. The overall pattern was almost identical to that found for memory load by number, which was also confirmed by a significant effect in the executive load and no reliable effect in the articulatory suppression condition.

Solution time and accuracy per answer component differed. Performance was best for the unit outcome and worst for the ten and hundred outcome. Accuracy per component was only moderated by value in such a way that more errors occurred on the outcome of the hundreds than of the thousands when the value of the carry was larger than one. Solution times per component depended on both number and value of the carries. For both number and value, the outcomes of the hundreds took more time with more carries or with larger carry values. With respect to value, the accuracy of tens, hundreds, and thousands was lower than of the units when the value was larger. This again shows that apart from number of carries, the value of carry also plays a prominent role in complex additions. It also appears that the effects of carries and their value increase as the calculation progresses, with the largest effects on the last position where a carry can occur: in the present problems, the hundreds.

GENERAL DISCUSSION

The problem addressed in the present article may be rephrased in four more elementary questions, namely (a) whether besides the number of carries the value of the carry is also an important determinant of calculation performance; (b) whether memory load, and an executive load in particular, augmented the effects of number and value of carries; (c) whether the two working-memory components studied, the phonological loop and the central executive, are differentially involved in performance variations due to number of carries and to value of the carries; and (d) whether specific effects occur in the different parts of the calculation problem. On the basis of an analysis of multidigit multisum arithmetic tasks in a vertical presentation format with incremental production of the answer, it was argued that participants would follow a step-by-step procedure to solve the problems. In other words, these problems may be viewed as a string of more simple arithmetic problems, which are not completely independent because occasionally a value of one component has to be carried to the next component. We argued that basically three factors would affect the difficulty of such problems. First, the difficulty goes hand in hand with the size of the outcome of each problem component. The larger this outcome, the more difficult this component would be. This derives from earlier findings and theorizing about the problem size effect in mental calculation (e.g., Ashcraft, 1992, 1995; Ashcraft & Battaglia, 1978; Butterworth et al., 2001; Campbell, 1995; Geary, 1996; Groen & Parkman, 1972). Second, difficulty also depends on the number of operations to be performed. When more digits have to be added, the size of the outcome being held constant, the problem will be more difficult, simply because there are more steps to be executed. If the outcome results in a value to be carried, this requires the execution of extra steps. Third, working memory mediates these operations. In producing the sum of a problem component, each step yields a result that must be maintained in working memory until the next step is complete, and the result should not be confused with the outcome of a previous sum. In other words, working memory must intervene to temporarily maintain a result (a storage aspect for which the phonological loop in the model of Baddeley & Hitch, 1974, might be thought responsible) and to prevent interference from previously encountered and maintained outcomes (blocking proactive interference for which the central executive would be the designated actor). The operation of these three factors was used to formulated predictions for the present experiments.

In the following paragraphs, we first summarize the findings and confront these with the predictions that we have formulated in the Introduction. The General Discussion is then concluded by elaborating on a number of issues that seem to play a crucial role in more complex arithmetic. In particular, we address the role of component problem size, interim results, and the role of working memory.

Summary of findings

In this paragraph, the present findings are summarized and are confronted with the predictions. This is organized around four themes: number and value of carries, their interaction with working-memory load, the role of phonological and executive working memory, and the problem components.

Number and value of the carry

First, the present study showed that value of the carry plays a role together with the number of carries. For one thing, the present study confirmed earlier findings by Fürst and Hitch (2000) and Noël et al. (2001) about the role of the *number of carries* in complex additions. The present study confirmed this by showing in Experiment 1 that problems with two carries were more difficult than problems with one carry in three-sum problems. Experiment 2 extended this finding to four-sum problems for a broader range of variation in the number of carries (1, 2, or 3). This effect can easily be explained on the basis of the augmented number of calculation steps in carry problems. With more carries, the number of additional calculation steps increases, and this results in slower performance. Assuming that each processing step is associated with a fixed probability of the occurrence of a calculation or a retrieval error, it also follows that when there are more processing steps, the probability of an error increases.

Interestingly, the present study also varied the *value of the carry* (values 1 and 2 in Experiment 1 and values 1, 2, or 3 in Experiment 2) and found that this variable also affects problem difficulty. Problems with carries of value 1 were faster and more correct than problems with carries of value 2. Problems with carries of value 2 were more accurate than problems with carries of value 3. Two elements may contribute to these effects. First, since in problems with carries one or more of the problem components will result in a larger outcome, problem size will, at least locally, increase with the value of the carry. As already stressed, problem solving is slower and more error prone when the outcome is larger. Second, when there is a carry, the value to be added to the next problem component must be kept in temporary storage, and then one should not forget to use this retained value in the calculation of the sum of the next component. Once again, a larger number of steps results in slower overall execution, and each step increases the overall probability of an error. In sum, larger values of the carry are on average related to a larger problem size and a greater number of processing steps. Both of these increase solution time and decrease accuracy.

Apart from main effects of number and value of the carries, the present findings also showed that with respect to solution time, the two effects interacted, while no such interaction was evident in accuracy. As was already argued in the Introduction, solution time increases with each additional carry, and since the execution time of each individual carry increases with the value of the carry, the combination yields a multiplicative effect (n * the individual time increment) of number and value of the carries. For accuracy, number of carries and value of the carries each increase the probability of an error to occur, and this results in an additive relationship.

Summarizing this part of the results, the present study shows that solution time of complex arithmetic sums depends on the number of carries, the value of the carries, and their interaction. Proportion of correct answers, in contrast, only depends on the number and the value of the carries. Two problem features seem to contribute to these effects, namely the number of processing steps and the size of the interim outcomes. The question remains, however, as to which mechanisms underlie these effects. Regarding the effect of number of carries, it is clear that working memory mediates these effects. In particular, it seems plausible that the central executive is involved in controlling the sequence of steps to be performed and to keep track of the progress. When there are carries, the procedure becomes more complex. This additional step must not only be planned, but care must also be taken to execute it, possibly several times. In fact, the role of the central executive entails additional control of the procedural sequence and monitoring of conflicts that may arise between the tendency to execute the no-carry sequence and the tendency to execute the planned sequence containing carries.

For the value of the carry, several mechanisms may come into play. On the one hand, when there is a carry, irrespective of its value, the number of processing steps increases, and this involves the mechanism described in the previous paragraph. On the other hand, the value of the carry must be maintained in temporary storage for later processing in the next part of the problem. Difficulties in maintaining this trace may result in slower as well as more error-prone processing, and these difficulties may be larger for larger values. Indeed, the error analysis suggests that more errors are made when the value is larger, but it also shows that the size of the error becomes larger with larger values of the carry, which seems to be due to either forgetting to add the carry or to adding too small a value, which suggests forgetting of the exact value of the carry. Trace decay would yield a simple, but probably untenable, explanation since there is no a priori reason why values of 3 would decay more easily than values of 2. A second possibility concerns interference from previous events. It may be easier to avoid interference or to recover from interference with values of 1 than with higher values. In fact, this may be coupled to effects of practice. We are more used to performing additions that require only a 1 to be carried. Practice effects have already been cited to explain the difficulty caused by the number of carry operations (Fürst & Hitch, 2000), but we believe it is possible that practice does not only account for the number, but also for the value to be carried. People are more used to performing additions with smaller carry values than with larger carry values. Moreover, when confronted with a series of calculations as in the present experimental sessions, the value carried in a previous problem may interfere with the value to be processed in the present one. Again, there is no a priori reason why larger values should be more interference prone than smaller values, but because the operations tend to be more difficult when the value is larger, it is possible that the temporary memory trace of a larger value is less resistant to interference. This may be due to a poorer maintenance of the trace in the competition with the calculation of the sum, as this competition is larger for larger outcomes, which are associated with larger values of the carry.

Load and number of carries

In some previously published research, it has been shown that the effect of the number of carries in an addition problem is augmented under a working-memory load and in particular under an executive load (e.g., Ashcraft & Kirk, 2001; Fürst & Hitch, 2000), whereas other studies failed to find a clear interaction of an executive load and number of carries (e.g., Logie et al., 1994). Together with the study of Fürst and Hitch (2000), both experiments of the present study show that a working-memory load adversely affected performance and that this effect was augmented in problems with multiple carries. While the data of the first experiment were not completely clear about this interaction, in the second experiment the accuracy data were sensitive enough to detect it. Decomposition of this interaction showed that this interaction completely bears on the presence of an executive load. These findings are consistent with previous research that also found that the central executive contributes to handling carries (e.g., Fürst & Hitch, 2000; Seitz & Schumann-Hengsteler, 2002), but they do not confirm previous findings that found a role for phonological effects in carrying (e.g., Ashcraft & Kirk, 2001; Noël et al., 2001; Seitz & Schumann-Hengsteler, 2002), and research that found no role for working memory in carrying (Logie et al., 1994).

The interaction between number of carries and executive load can be explained by referring to a competition between two task sequences, one without and one with an execution of a carry. When there is a carry, the value is temporarily maintained, and it must be remembered (a) to perform the carry and (b) to clear the memory after performing the carry. The working-memory load incurred by these processes is rather small. However, in the context of a sequence of carries interference may arise because a previously stored value is added again (which is rather rare) or because a to-be-stored value is not maintained because one forgets to maintain this value (which is a rather frequent mistake in the present data). The interference arising in these contexts will affect accuracy of performance rather than the speed of performance, and this could also explain why the interaction of load and number was only reliable in the accuracy data.

Load and value of the carries

As already stipulated, the present findings show that the value of the carry contributes to the difficulty of multisum problems. In the accuracy data, but not in the solution times, of both experiments, this effect was augmented by the presence of a memory load. Decomposition of the interactions has shown that this interaction was only present with an executive load. Although it seems obvious that the value to be carried is maintained in the phonological loop, the load imposed by this storage does not seem to be big, as the trace is not disturbed by the presence of articulatory suppression. An executive load, on the contrary, seems to impair calculation with carries and the more so when the value of the carry is larger. Once more, this can be explained in terms of control for proactive interference due to intrusion of previously used values of the carry. In view of the fact that thus far the value of the carries in (complex) mental arithmetic has not received much attention, the conclusion that maintenance of the value of the carries seems to rely on executive control—more specifically, interference control—is an important one that deserves follow-up in future research.

Differentiating the role of working-memory components

The present results show that both the executive memory component and the phonological loop play a role in carrying out complex additions. The effect due to the phonological load was, however, only present as a general effect; it did not interact with either number of carries or value of the carry. We believe that the phonological loop plays a role in solving these problems, but the sensitivity of these kinds of experiment is probably not strong enough to detect the effect imposed by a phonological load, or alternatively the involvement of the phonological loop is too restricted to be detected in a selective interference study. In particular, in the present design, the role of the phonological loop may have been minimized by the requirement to type each outcome part as it became available. This obviated the need for storing interim results. A consequence of this may be that the present findings about the effects of a phonological load are not comparable to the effects reported in other studies.

The effect of the executive loads, on the contrary, was clearly present in the solution time and the accuracy data of both experiments as a general effect and in the accuracy data of Experiment 2 in the interaction with number of carries and value of the carry. As we have already extensively discussed in the previous paragraphs, all these findings clearly point to the involvement of the central executive to control the calculation process and especially to control the execution of the calculation in the face of possible intrusions or proactive interference. There is little doubt that these control processes must act in interplay with the contents of the phonological loop, but as already mentioned, the present study does not provide any direct support for this hypothesis. Nevertheless, a cooperation of the central executive with the phonological loop would seem more plausible (see, e.g., Emerson & Miyake, 2003; Liefooghe, Vandierendonck, Muyllaert, & Van Neste, 2005; Miyake, Emerson, Padilla, & Ahn, 2004), but has to wait for more direct supporting data.

Problem components

A further novel feature of the present study concerned the focus on the answer components as they were registered online in an incremental response. On the basis of an analysis of the operations involved in the different components, it was predicted that the solution time and accuracy of the components would depend on their position in the sequence as this position is related to the amount and the extent of processing involved at each position. In this vein, it was predicted that the answer for the units part would be faster and more accurate than that for the other parts and that the final position (hundreds in Experiment 1; thousands in Experiment 2) would also tend to be easier than the middle positions. These predictions were confirmed both for solution time (where the main effect was only marginally significant in Experiment 1) and for accuracy.

No predictions were formulated about the interaction of the components with the other effects in the design. The main reason for this restraint is that so many variations may be taking place at each individual calculation that this would make sense only in a much larger study based on a completely controlled set of additions. Indeed, the present data are quite variable in this respect over the two dependent measures in the two experiments. Nevertheless, the level of detail achieved with this online registration and analysis procedure seems to be very promising in the collection of more detailed data about mental arithmetic performance.

Implications for models of mental arithmetic

The main findings of the present study seem to be as follows: (a) It confirms previous results that show that number of carries determines problem difficulty and that this is quite likely mediated by working memory's executive subsystem; (b) the present study shows furthermore, that the value of the carries also makes an important contribution to problem difficulty and that this also relies on control processes of the central executive; (c) with respect to solution time, the effect of number of carries augments the effect of value of the carries; (d) even though general effects of both phonological and executive loads were observed, only the executive loads modulate the effects of number and value of the carries; (e) the different steps in the production of the answer are not equivalent and are dependent on the number of operations and the difficulty of the operations performed in each part of the problem.

The theoretical analysis for the present study was based on the hypothesis that participants solve these complex additions by partitioning the calculations into a sequence of smaller one-digit sums in such a way that the load on working memory is kept minimal. This analysis resulted in a set of predictions that were corroborated in the two experiments of this study. A key notion in this theoretical analysis was that the difficulty of each subproblem depends on two elements, namely (sub)problem size and number of operations in the subproblem.

Number and value of the carries seem to be strongly correlated to a measure of problem size based on the sums of all subproblems. The question may be raised as to whether it would not be simpler to consider all the effects in terms of problem size rather than in terms of number and value of the carries. First, although the effects are related to problem size, this factor covers only part of the effects. Apart from problem size, our analysis also included the number of operations performed and the working-memory basis of these operations. The findings indicate that all three elements should be taken into account. Considering only problem size as the explanatory factor will undoubtedly result in a deficient explanation. Second, problem size is merely a characteristic of the problems that does not clarify for which reason problems with a larger size would be more difficult. Specification of the effects in terms of the number of carries and the value of the carries does seem to be more helpful in that these notions are directly linked to operations performed in calculation. That there is a relationship between number and value of the carries on the one hand and summed problem size on the other hand is nevertheless useful as it clarifies that one of the factors that makes an addition more difficult is the size of the subproblems.

In explanations referring to working-memory mediation of complex mental arithmetic performance, the role of the maintenance of interim results is often stressed. The theoretical analysis on which the present study was based deviates from such a view in important ways. First, by looking at incremental solutions of additions, there is no need to maintain the interim results until the complete answer can be emitted. By relaxing this artificial memory requirement, it becomes possible to look at how working memory is necessarily involved in these additions. If people indeed decompose the problem in a series of successively solved subproblems—and the present findings are consistent with such a view—then the need for maintenance of interim results becomes quite small. Indeed, the theoretical analysis developed in the Introduction suggests that carries have to be maintained for a very brief period, while the response for the current subproblem is being prepared. Immediately thereafter, calculation can continue, and the value carried can be taken as the first value to be added in the next subproblem. This way, at any one time only one value must be kept in memory—namely, the last obtained sum. In addition to that, working memory is needed to keep track of progress in order to make sure that every digit is entered in the calculation once and not more than once.

Implicit in this description of processes is the notion of sequential processing. More specifically, it is suggested that calculation stops while the response is being prepared. This is a possible reading of the description. Our basic assumption is that only one calculation step is performed at one time. Having obtained the result for one column in the addition problem, if the outcome exceeds 10, the value of the tens must be carried, and the value of the units must be produced as the answer. Minimally, during this operation, the value of the carry must be maintained. Presently, it is not clear how much effort is needed to prepare the answer to be emitted. To the extent that this entails a selection of the appropriate response, it may also be expected that further calculation is interrupted during this period (cf. PRP effect; Pashler, 1984).

Working-memory involvement in complex additions goes beyond keeping track of interim results and sequencing of the solution steps. Every operation performed may call on working-memory resources. When the addition problem is decomposed in simpler problems, a sequence of simple additions must be performed. Each simple addition consists either of a retrieval of the outcome from long-term memory or of a transformation of the problem to a problem to which the outcome is known. In both cases, retrieval of an outcome from long-term memory is required. Although some theorists seem to assume that retrieval is an automatic process, it is more likely that memory retrieval requires cognitive control (Barrouillet, Bernardin, & Camos, 2004; Deschuyteneer & Vandierendonck, 2005a, 2005b; Szmalec et al., 2005). Therefore, in addition to the maintenance of the interim results and the sequencing and control of the successive calculation steps, working memory is also involved in retrieval of the outcome. Of all these operations, only the latter one is also at the basis of working-memory engagement in simple mental arithmetic.

The results of the present study are completely based on complex additions. The idea that such problems are partitioned into smaller problems could also be valid for other complex arithmetic operations such as multiplications, subtractions, and divisions. Further research with these operations is needed, though, to test the generality of the present view of working-memory mediation in complex arithmetic.

Original manuscript received 17 May 2005
Accepted revision received 11 April 2006
First published online 20 July 2006

REFERENCES

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*, 75–106.

Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: A review and summary of new directions. *Mathematical Cognition, 1*, 3–34.

Ashcraft, M. H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 527–538.

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130*, 224–237.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A*, 5–28.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133*, 83–100.

Butterworth, B., Zorzi, M., Girelli, L., & Jonckheere, A. R. (2001). Storage and retrieval of addition facts: The role of number comparison. *Quarterly Journal of Experimental Psychology, 54A*, 1005–1029.

Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition, 1*, 121–164.

De Rammelaere, S., Stuyven, E., & Vandierendonck, A. (1999). The contribution of working memory resources in the verification of simple mental arithmetic sums. *Psychological Research, 62*, 72–77.

De Rammelaere, S., Stuyven, E., & Vandierendonck, A. (2001). Verifying simple arithmetic sums and products: Are the phonological loop and the central executive involved? *Memory & Cognition, 29*, 267–274.

De Rammelaere, S., & Vandierendonck, A. (2001). Are executive processes used to solve simple mental arithmetic production tasks? *Current Psychology Letters: Behaviour, Brain & Cognition, 2*, 79–89.

Deschuyteneer, M., & Vandierendonck, A. (2005a). Are "input monitoring" and "response selection" involved in solving simple mental additions. *European Journal of Cognitive Psychology, 17*, 343–370.

Deschuyteneer, M., & Vandierendonck, A. (2005b). The role of "response selection" and "input monitoring" in solving simple arithmetical products. *Memory & Cognition, 33*, 1472–1483.

Deschuyteneer, M., Vandierendonck, A., & Muyllaert, I. (in press). Does solution of mental arithmetic problems such as 2 + 6 and 3 * 8 rely on the process of "memory updating"? *Experimental Psychology*.

DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*, 353–386.

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language, 48*, 148–168.

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition, 28*, 774–782.

Geary, D. C. (1996). The problem-size effect in mental addition: Developmental and cross-national trends. *Mathematical Cognition, 2*, 63–93.

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review, 79*, 329–343.

Heathcote, D. (1994). The role of visual–spatial working memory in the mental addition of multi-digit addends. *Current Psychology of Cognition, 13*, 207–245.

Hitch, G. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology, 10*, 302–323.

Imbo, I., Vandierendonck, A., & Vergauwe, E. (in press). The role of working memory in carrying and borrowing. *Psychological Research*.

Lemaire, P., Abdi, H., & Fayol, M. (1996). The role of working memory resources in simple cognitive arithmetic. *European Journal of Cognitive Psychology, 8*, 73–103.

Liefooghe, B., Vandierendonck, A., Muyllaert, I., & Van Neste, W. (2005). The phonological loop in task alternation and task repetition. *Memory, 13*, 550–560.

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition, 22*, 395–410.

Lorch, R. F., Jr., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 149–157.

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. C. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115*, 123–142.

Noël, M.-P., Désert, M., Aubrun, A., & Seron, X. (2001). Involvement of short-term memory in complex mental calculation. *Memory & Cognition, 29*, 34–42.

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 358–377.

Roth, P. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology, 47*, 537–560.

Seitz, K., & Schumann-Hengsteler, R. (2002). Phonological loop and central executive processes in mental addition and multiplication. *Psychologische Beiträge, 44*, 275–302.

Szmalec, A., Vandierendonck, A., & Kemps, E. (2005). Response selection involves executive control: Evidence from the selective interference paradigm. *Memory & Cognition, 33*, 531–541.

Trbovich, P. L., & LeFevre, J. A. (2003). Phonological and visual working memory in mental addition. *Memory & Cognition, 31*, 738–745.